

# Newcastle University e-prints

---

**Date deposited:** 30<sup>th</sup> March 2010

**Version of file:** Author final , book chapter

**Peer Review Status:** Peer Reviewed

## Citation for published item:

Wright C. [The role of Working Memory in the development of L2 grammatical proficiency](#). In: Benati, AG, ed. *Issues in Second Language Proficiency*. London: Continuum, 2009, pp. 45-60.

## Further information on publisher website:

<http://www.continuumbooks.com/>

## Publishers copyright statement:

The definitive version of this chapter, published by Continuum Books, 2009.

The definitive version should always be used when citing.

---

## Use Policy:

The full-text may be used and/or reproduced and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not for profit purposes provided that:

- A full bibliographic reference is made to the original source
- A link is made to the metadata record in Newcastle E-prints
- The full text is not changed in any way.

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

<p><b>Robinson Library, University of Newcastle upon Tyne, Newcastle upon Tyne.</b> <b>NE1 7RU. Tel. 0191 222 6000</b></p>
--

## Chapter 4

# The Role of Working Memory in the Development of L2 Grammatical Proficiency

Clare Wright

*Newcastle University, United Kingdom*

### Abstract

This study investigates the correlation between Working Memory (WM) capacity and individual differences in the development of L2 grammatical proficiency in an immersion setting. Adult Chinese speakers of English were tested over a ten-month period for acquisition of English question forms using timed oral and written tasks, and a battery of WM tasks. Significant differences were found between individuals' linguistic development, and between task mode (oral vs. written), and question type (subject vs. object). Positive (though non-significant) correlations were found between linguistic and WM scores, supporting the claim (Miyake & Friedman, 1998) that WM plays a role in L2 development.

### Introduction

The study discussed here examines the potential interaction between WM capacity and individual variation in adult acquisition of a second language (L2), with specific reference to acquisition of English wh-questions (wh-movement) by instructed Chinese speakers of English.

Wh-movement has been long identified as an area of individual variation in morpho-syntactic proficiency in adult L2 acquisition (Johnson & Newport, 1991). WM (Baddeley, 1986, 2000) has been claimed to be 'key' to understanding such L2 variation (Miyake & Friedman, 1998), following previous investigations into connections between memory, processing and L2 acquisition (see, for example, Brown & Hulme, 1992; Lado, 1965). Robust correlations have been found between WM capacity and acquisition of L2 vocabulary, oral fluency and reading skills (Fortkamp, 1999; Harrington & Sawyer, 1992; Service, 1992). However, longitudinal research to test this claim for morphosyntactic acquisition, especially when tested orally, remains sparse, with virtually none from a generative perspective (Juffs, 2004).

This study explores some of the assumptions and methodology in current research by investigating if WM capacity is a significant factor in individual variation in rate of acquisition for CSE in an immersion environment. Two semi-longitudinal studies were carried out assessing acquisition of wh-question formation by advanced CSE during study-abroad periods in the United Kingdom. Eleven participants took part in a preliminary exploratory study, presented in detail here, using a battery of innovative linguistic tests (oral and written) and WM tests. Individuals' rates of acquisition differed significantly, and patterns of asymmetry were also found between task mode (oral vs. written), and question type (subject vs. object). Positive (though non-significant) correlations between rates of L2 acquisition and WM provided some support for the research hypotheses tested here, although a number of methodological issues arose, which are being addressed in a second larger study currently being undertaken.

## Theoretical and Empirical Background

English wh-question formation was identified as a source of wide individual variation in acquisition, particularly for CSE, even at advanced levels and after immersion in English (Han, 2004; Johnson & Newport, 1991; Schachter & Yip, 1990; Wright, 2006). Wh-questions require syntactic movement while Chinese lacks overt wh-movement; acquisition of the L2 linguistic knowledge to produce accurate wh-questions is seen as late acquired (Pienemann, 1998), especially for long-distance movement such as 'What did John say Mary wanted?' or constrained by subjacency, such as \*'What did the book about please Mary?'

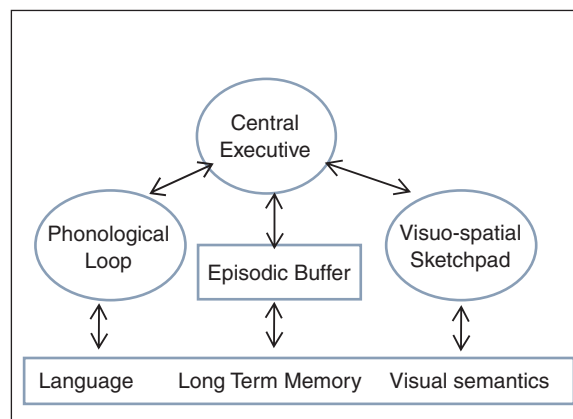
Standard generative accounts of whether CSE acquire wh-movement remain inconclusive (see, e.g. White & Genesee, 1996; Hawkins & Chan, 1997; White, 2003; Schwartz, Ma & Kim, 2008), while other accounts of wh-movement suggest that general cognitive processes explain such individual variation (Clahsen & Felser, 2006; Johnson & Newport, 1991; McDonald, 2006). The question of how input triggers acquisition of the L2 (Carroll, 2001; Sakas & Fodor, 2001; Schwartz, 1993) also remains debated. CSE have commonly been taught in an input-poor learning environment, where a traditional emphasis on explicit and written linguistic knowledge can result in wide variation between oral and written proficiency (Gu, 2003; R. Ellis, 1994), although research suggests that CSE can reach native-like competence even without exposure to native language immersion (White & Juffs, 1998). Focusing on acquisition of wh-movement should therefore provide insight into how instructed L2ers access explicitly learned structures (short-distance wh-movement, long-distance wh-movement), and structures assumed to be implicitly learned (subjacency).

The 'coalitionist' model proposed by Herschensohn (2000) is used here as a construct which allows some interface between the generative and non-generative research paradigms referred to above. In this model, Herschensohn

argues ‘that the L2er uses a coalition of resources’ including a UG template, L1 transfer, primary linguistic data and ‘instructional bootstrapping’ (ibid: 220). She discusses how the language user is able to draw on both implicit and explicit knowledge about language ‘presumably located outside the language module in the knowledge base’ (ibid: 184–85), and accessible either, in simplified terms, via implicit procedural memory, or explicit declarative memory (Paradis, 2004; Ullman, 2001).

The assumption drawn here is that L2 users with little naturalistic L2 input, but using ‘instructional bootstrapping’, will initially store L2 linguistic knowledge of morphology and syntax primarily as consciously accessible or explicit knowledge. After sufficient exposure to primary linguistic data, implicit knowledge develops, subject to the UG template. Until the L2 user can utilize the quicker, more efficient but non-accessible implicit system, it is assumed that conscious access of explicit knowledge is key for the L2 user. Therefore WM, the temporary ‘workspace’ for conscious attention to complex tasks, will also be key in efficient retrieving or inhibiting existing explicit knowledge and processing novel information (Smith & Kosslyn, 2007, 247).

Much of the research on WM in native language and L2 acquisition has been based on versions of Baddeley and Hitch’s (1974) multi-component model. The latest model (Baddeley, 2000, 2003) posits domain-specific temporary storage via the phonological loop and the visuo-spatial sketch-pad. Domain-general attentional control and processing efficiency is through a central executive, with an episodic buffer allowing domain-general storage for more than the standard 1-2 seconds (see Figure 4.1). The episodic buffer, a new element (Baddeley, 2000), is designed to explain how novel and retrieved information can be combined and maintained, for example allowing a prose passage of around 90 seconds to be retained and repeated accurately.



**FIGURE 4.1** Baddeley’s (2000) Multicomponent Model of Working Memory

WM is seen as capacity-constrained: as storage capacity reaches or exceeds its limit, processing efficiency is reduced, so greater storage capacity allows greater processing efficiency. Initial research findings using WM tests for phonological loop storage and central executive efficiency have found a robust correlation between WM and native-language (L1) proficiency (Gathercole, 2006; Gathercole & Baddeley, 1993). This robustness extends to certain aspects of L2 proficiency: for vocabulary acquisition, reading comprehension, resolving syntactic ambiguities and oral fluency (Baddeley, Gathercole & Papagno, 1998; Ellis & Sinclair, 1996; Fortkamp, 1999; Harrington & Sawyer, 1992; Miyake & Friedman, 1998; Osaka & Osaka, 1992; Service, 1992).

The evidence leads to three assumptions underpinning the current study of ways in which WM may be key to L2 acquisition even at an advanced level. The first relates to the role of WM in attentional control, where WM acts as a kind of 'bottle-neck' (Emerson, Miyake & Rettinger, 1999) through which L2 linguistic operations have to pass – the more novel the sound, or the more complex the task, the more significant the capacity of the 'bottle-neck'. For L2 users, producing target-like (or accurate) morphosyntax under pressure, such as in spontaneous speech or timed grammaticality tasks, requires conscious control over accessing explicit L2 knowledge and inhibiting L1 language patterns.

The second assumption is that the more difficult the morphosyntax (such as for long-distance wh-movement), the greater the effort in processing accurate forms. This assumption is supported by research into the role of WM in native-language complex syntax, such as assigning relative-clause reference (Miyake, Carpenter & Just, 1994) and subordination and adverbial use (Fry, 2002). Furthermore, processing difficulties are hypothesized as an explanation for individual variation in L2 (Cook, 1997; Service, Simola, Metsanheimo & Maury, 2002) even in generative accounts such as White and Juffs (1998), who explained evidence of variation in native-like L2 acquisition as 'implicit competence processed more slowly' (ibid: 127).

The third assumption concerns linguistic development in an immersion setting: that WM may be hypothesised to play a part in the processing of primary linguistic input, through processing novel acoustic information via the phonological loop, which may be particularly important when the amount and type of input changes in an immersion setting. For example, following this assumption, greater WM capacity would be predicted to facilitate quicker transition through developmental stages, or facilitate greater accuracy. This assumption follows psycholinguistic research into the major role played by noticing in converting input to intake (Carroll, 2001; Sagarra, 2007; Schmidt, 1990; VanPatten, 1996; VanPatten, 2005).

There is however increasing concern about the reliability and validity of commonly used WM tasks (Juffs & Rodriguez, 2006; Yoshimura, 2001), which casts doubt on the role WM can play across general L2 proficiency. The most

common tasks consist of storage-only measurements (such as digit span, word span, non-word repetition) or storage plus processing measurements (such as reading span, speaking span, elicited imitation). Non-word repetition has been one of the most robust tools for correlations with L1 and L2 vocabulary learning (see Gathercole, 2006 for a recent overview). Daneman and Carpenter's (1980) Reading Span Measure (RSM) and Daneman and Green's (1986) Speaking Span Measure (SSM) have also been widely used (Harrington & Sawyer, 1992; Osaka & Osaka, 1992; Sagarra, 2000; Yoshimura, 2001 for RSM, Fortkamp, 1999; Mizera, 2006 for SSM). However, early suggestions that these storage plus processing tasks would yield strong correlations with general L2 proficiency have not been borne out in further research, and current thinking is that WM is highly task-specific: for instance, RSM may only correlate with L2 reading proficiency (Yoshimura, 2001). Other research even contradicts some of the findings suggested above (Mizera, 2006 found no relation between oral fluency and WM, unlike Fortkamp, 1999).

There is also some debate over Baddeley's multi-component model used in the L2 studies cited above (Andrade, 2001; Caplan & Waters, 1999; Cowan, 1999; Miyake & Shah, 1999). Some models of language learning (Clahsen & Felser, 2006; Craik & Lockhart, 1972; Shallow Structure Hypothesis) identify 'depth of processing' as the key to successful transfer of linguistic knowledge between long-term and short-term memory. Finally conceptual issues remain unresolved such as how WM might interface with different memory models for language acquisition (N. Ellis, 2005; Jackendoff, 1997; Schmidt, 1994; Skehan, 1998), and how WM is separable from general processing constraints in L2 (Clahsen & Felser, 2006; Cook, 1997; Juffs, 2004; Sagarra, 2007). Nevertheless, Baddeley's model remains the most widely used for researching WM in language acquisition and is thus the basis for this study.

## Methodology

### Study Design and Participants

- i. To investigate the conceptual and methodological issues raised above, a semi-longitudinal study was designed to test for positive correlations between WM and L2 variation in acquisition by adult Chinese speakers of English during a study-abroad period. Two research hypotheses were addressed:
- ii. Immersion facilitates acquisition of wh-movement; WM correlates with development of wh-movement in an immersion environment.

Linguistic data was collected at Time 1 (within two months of arrival) and at Time 2 (after nine to ten months' immersion). WM data was collected at Time

1 and at Time 2. Since a directional correlation was assumed between WM and linguistic development, only WM scores from Time 1 were used in the analysis reported here.

Eleven advanced adult speakers of English, with Mandarin Chinese as L1, were recruited from a cohort of newly arrived postgraduates at UK universities<sup>1</sup>. All were instructed learners with no previous immersion exposure, with a minimum IELTS score of 5.5 in the previous four months. The group consisted of three participants from Taiwan, and eight from Mainland China; there were four males and seven females. Bio-data on learning background and exposure to input were gathered, to test for inter-learner variation in exposure to English prior to immersion (Dornyei, 2003), and no significant differences were found.

### **Linguistic Data Collection**

Linguistic data were collected using two time-constrained tests of oral and written proficiency in question formation: a guided oral question and answer 2-way gap fill task (Task 1), and a written grammaticality judgement task (Task 2), adapted from White and Juffs (1998).

#### ***Task 1: Oral Production (OP)***

Task 1 measured question forms produced in a seven-minute dialogue to complete a picture of a party scene. Following a commonly accepted hierarchy of acquisition (e.g., Pienemann, 1998), question forms produced by the participants were divided into two groups: Group 1 (formulaic chunks, intonation only, question word fronting without head movement and copula fronting), and Group 2 (head movement and 'do'-support, cancelling inversion in embedded clauses), such as 'Will he come later?', 'What did the girl eat?', 'Can you tell me when the boy arrived?'.

Target-like (native-like) production of questions from Group 2 were taken to imply acquisition of wh-movement. The total number of utterances was also measured for evidence of task avoidance (Schachter, 1974). The total number of Group 2 questions was then divided by the total number of utterances during the seven-minute test to produce an OP score.

#### ***Task 2: Grammaticality Judgement (GJ)***

A written task (Task 2) was also used in view of the difficulty in analysing oral data of untangling 'performance noise' from underlying competence (Murphy, 1997). Task 2, based on White and Juffs (1998), asked for graded judgements (using a Likert scale of -3 to +3) of grammatical acceptability on complex

question forms, derived from the party scene used in the first task. Scoring was calculated for native-like accuracy on 22 tokens of subject and object long-distance movement questions and subjacency violations (10 ungrammatical and 12 grammatical tokens). Examples of ungrammatical tokens are given below:

- |  |                     |
|--|---------------------|
| (1) *Who did Tom expect give the present?        | <i>Subject</i>      |
| (2) *What did John know did Ann like?            | <i>Object</i>       |
| (3) *What did Tom bring a present after he sent? | <i>Subjacency</i> * |

### Working Memory Data Collection

In line with current best practice in testing WM (see, e.g., Conway et al. 2005), a battery of tasks was used: a non-verbal task (Digits Back), and two innovative verbal tasks, designed for this study (Story Recall, and a combined Word Span and Sentence Span task).

#### *Task A: Digits Back (DB L1, DB L2)*

The first task, Digits Back, was chosen as being widely used (Waters & Caplan, 2003), and therefore providing a reliable benchmark, easy to administer, and unrelated to linguistic proficiency (since advanced learners would all be familiar with the English digit names). It would also provide a cross-linguistic comparison between individuals' scores in Mandarin and English, in the light of research showing differences between Digit Span scores in Mandarin and other languages (e.g. Chincotta and Underwood, 1997).<sup>2</sup> This would shed further light into investigations as to how far WM is language-independent (Osaka & Osaka, 1992) or affected by differences in L2 processing (Cook, 1997; Service et al., 2002). Participants heard sets of numbers, increasing in length from four to seven, read out at a rate of one digit per second (two strings per set) first in English and later in Mandarin. After each string, participants repeated the numbers in reverse order. Scoring was calculated following an 'all or nothing' score, using the length of the set where two strings were last recalled correctly (Conway et al. 2005: 774).

#### *Task B: Story Recall (SR L1, SR L2)*

Task B was created to address the issue of how to test newer models of the WM construct, in particular Baddeley's episodic buffer for which virtually no research on L2 WM has yet been published.<sup>3</sup>

The Story Recall task devised for this study was based on standard psychology tests (see, e.g., Coughlan & Hollows, 1985), which measures the accuracy in



recalling prose passages of over 30 seconds in length. This test had been identified as correlating with use of complex syntax such as subordinate clauses and adverbial phrases in native language research (Fry, 2002). The purpose of using the task in this study was to assess whether it was also a reliable, valid means of testing WM in L2, and if WM as measured by such a test correlated with the use of complex question formation.

The original Coughlan and Hollows test was adapted and translated into Mandarin (lasting 54 seconds). A different story in English (with similar schematic structure) was devised by the researcher (lasting 33 seconds). The length of the L2 was shortened to avoid possible 'floor effects' due to task difficulty (Harrington & Sawyer, 1992: 28). Two bilingual raters worked with the researcher in scoring the Mandarin data to ensure scoring reliability. Scoring for the task (SR) was out of 50 for accurate recall of morpho-syntactic and semantic elements.

### ***Task C: Word Span and Sentence Span (WS, SS)***

Task C was adapted from Daneman and Carpenter's (1980) Listening Span task. It was devised to examine the offset between storage and processing through the combined testing of phonological short-term storage (Word Span) and central executive capacity (Sentence Span), using a direction-based task.

Pairs of directions using the words 'left', 'right', 'up' or 'down' were created, increasing by one in each pair from five to twelve words. An example pair of sentences is given in (4) below.

- (4) (i) Walk **up** the street until the lights. (length: 7 words)  
 (ii) Take the second turn on the **left**.

Participants heard, then repeated, both strings; the repetition provided a secondary processing task. Scoring, like the Digits Back test, was a 'quasi-absolute' score given here in percentage form, indexing two measures. The score for Sentence Span (SS) was the longest sentence length (measured in number of words out of 12) when direction words were correctly recalled. The score for Word Span (WS) was the longest sentence length when all the words in the sentence were correctly recalled. Due to limitations of time, this task was only done in English (L2).

## **Results**

All the scores for linguistic and WM data were encoded using SPSS, and converted to percentage scores shown here, for ease of comparison across all tasks. Non-parametric statistical tests were used, given the small sample size ( $n = 11$ ).

The minimum, maximum and mean scores on linguistic and WM tasks at Time 1 and linguistic tasks at Time 2 are shown in Tables 4.1 and 4.2. Wilcoxon signed rank tests of difference were performed to check for significant differences between the linguistic scores at the two different test times, and for inter-language differences in the WM tasks in Mandarin and English. Tests of difference on the linguistic tasks showed that the two OP scores were not significantly different ( $p > .05$ ), but that the GJ scores were significantly different ( $p = .007$ ). Tests of difference for the WM tasks comparing L1 and L2 showed that the DB scores were not significantly different ( $p > .01$ ) but the SR scores were significantly different ( $p = .005$ ).

Individuals' linguistic scores at Time 2 were then correlated with WM across all tasks, using Spearman's rho to test for evidence of positive correlation in line with the second research hypothesis that WM would affect linguistic development during immersion. The correlations were not significant for any of the tests, although there were moderate positive correlations between SS and OP ( $r = .32$ ) and GJ ( $r = .28$ ). The results are shown in Table 4.3.

The lack of significant findings initially appear to provide poor support for the research hypotheses tested here. However, on closer examination of the results, it was clear that using a simple group mean score at Time 2 obscured unexpected and more interesting individual variation in patterns of linguistic development between Time 1 and Time 2. The data were therefore assessed to show individuals' linguistic development over the time of the study, that is, how much individuals' linguistic scores showed a change (either negative or positive). Individuals' changes in GJ scores (Task 2) were further analysed

**Table 4.1** Linguistic data scores

Linguistic Scores by Task	Minimum	Maximum	Mean
OP Time 1	0	27	9
OP Time 2	8	31	19
GJ Time 1	18	46	37
GJ Time 2	36	68	50

**Table 4.2** WM data scores

Working Memory Scores	Minimum	Maximum	Mean
DB L1	77	100	91
DB L2	57	86	79
SR L1	64	96	81
SR L2	14	62	44
Word Span	42	83	61
Sentence Span	42	100	75

**Table 4.3** Correlation of linguistic and WM scores

	<b>DBL1</b>	<b>DBL2</b>	<b>SRL1</b>	<b>SRL2</b>	<b>WS</b>	<b>SS</b>
OP Time 2	-0.02	-0.09	0.13	0.04	0.14	0.32
Sig.	0.96	0.80	0.72	0.91	0.68	0.34
GJ Time 2	0.28	0.23	0.12	0.19	0.25	0.28
Sig.	0.41	0.53	0.73	0.59	0.46	0.41

**Table 4.4** Variation in linguistic development

	<b>Minimum</b>	<b>Maximum</b>	<b>Mean</b>
Change in OP	-25	22	4
Change in GJ	-1	23	10
Change in objects (GJ)	-10	50	17
Change in subjects (GJ)	-30	50	5

**Table 4.5** Correlation of variation in linguistic development and WM scores

	<b>DB L1</b>	<b>SR L1</b>	<b>SR L2</b>	<b>WS</b>	<b>SS</b>
Change OP	0.43	0.33	0.14	-0.01	0.27
Sig.	0.19	0.35	0.69	0.98	0.42
Change obj	0.01	-0.13	-0.06	-0.37	-0.38
Sig.	0.97	0.71	0.87	0.26	0.25
Change sub	0.30	0.45	<b>0.60</b>	0.13	-0.05
Sig.	0.38	0.20	<b>0.05</b>	0.69	0.87

by question type, yielding a significant subject-object asymmetry ( $p = .003$ ). Results are shown in Table 4.4 (in percentage terms, as for the earlier results).

These scores were then correlated with WM scores to see more precisely if WM scores at Time 1 showed a positive correlation with the degree of change in both tasks, and by question type in Task 2. No significant results were found for the change in grammaticality judgements (GJ, Task 2): all correlations were negative except for Digits Back in L1 ( $r = .23$ ,  $p > .005$ ); in addition, no significant results were found for Digits Back in L2. Therefore, results for changes in GJ and DB L2 are not shown here. However, positive correlations were found for change in oral production (OP, Task 1), and asymmetric correlations were found with changes in subject and object grammaticality judgements (GJ, Task 2). These results are shown in Table 4.5.

There was one significant correlation between the change in subject accuracy and Story Recall in L2 ( $p = .05$ ), shown in Table 4.5 in bold.

## Discussion

There were no clear patterns of either linguistic development or of significant correlation between linguistic and WM tasks, so there is no robust support for the research hypotheses that immersion would facilitate acquisition and that WM would correlate with linguistic development during immersion.

The first general conclusion is that immersion for nine months was not long enough to trigger significant development across the whole group, as evident from the unexpected wide range of individual variation; indeed four out of eleven participants showed negative scores for their linguistic development in both oral and written tasks. However, the evidence of significant subject-object asymmetry (found also by Schachter & Yip, 1990 and White & Juffs, 1998) indicates that linguistic development may occur at different rates, not just by individual but by linguistic phenomenon. Further research using more fine-grained measures of linguistic development would be able to test this suggestion.

Second, the different WM task scores revealed little consistency; DB and SR did not correlate across language, suggesting they may be language-dependent (contrary to Osaka & Osaka, 1992) and there were no significant correlations between the tasks, suggesting that WM tasks are also task-dependent (as found by Yoshimura, 2001). However, the innovative nature of the verbal WM tasks (SR, WS and SS) may have made them difficult to compare with other WM studies using the standard Daneman and Carpenter (1980) test.

Third, the different WM tasks showed no consistent pattern of correlation with either oral or written linguistic data. The strongest positive correlations (although not significant) were found between the oral data (OP) and WM, ranging from  $r = -0.01$  to  $r = .43$ . Given that OP was tested in an online task, and WM is hypothesised to relate to managing complex online or pressured tasks (Baddeley, 2003, Sagarra, 2007), this finding was as expected. The strongest correlations were for WM tested in L1, thought to be a 'purer' measure of WM, not confounded by issues of L2 language proficiency (Service et al. 2002) and thus supporting the general assumption that WM capacity plays a role in L2. There was a significant strong correlation between SR L2 and change in subject question accuracy ( $r = .60$ ,  $p = 0.05$ ) in the offline written grammaticality judgement task (GJ), offering some evidence that WM capacity affects L2 even in less pressured tasks. However, another, linguistic, explanation could be that L2 proficiency at the GJ task may help L2 performance in the SR task. The subject-object asymmetry could be explained by differences in processing subject extraction compared with

objects, as suggested by White and Juffs (1998). Further more fine-grained data collection to include processing information (such as using a reaction time test) would allow this suggestion to be tested.

In addition to the points identified above, the study was also affected by a number of methodological issues. First, the small group size reduced the likelihood of significant findings; in addition, the drop out of participants meant that the original interview pairings for the oral task (Task 1) at Time 1 were not exactly replicated at Time 2; furthermore, limiting the scoring from Task 1 to only full lexical wh-questions produced very small numbers of raw data. Second, the design of the Word Span and Sentence Span, trying to measure both phonological storage and executive control in a single test, could have undermined the reliability of the test design. Additionally, the 'quasi-absolute' scoring system is no longer seen as the optimal scoring system (Conway et al. 2005), particularly for such small groups, whereas a 'partial' scoring system which takes account of more individual variation in results, is seen as more valid.

In order to address these issues, a second larger study has been devised, with more participants ( $n = 30$ ). The GJ task has been redesigned as a computer-based RT task (DMDX) to allow for greater precision on patterns of processing of different types of wh-movement, including subject and object extraction. The WM listening span task and scoring have been redesigned to fit more closely with existing WM methodology. The study is still ongoing, although initial analysis of the reaction time data from the first point of collection (Time 1) replicates the significant difference in accuracy on subject-object judgements ( $p = .004$ ) found in the first study. Data from Time 2 is anticipated to show how far WM may play a role in individual variation in improvements in oral output, and in faster and more accurate RT scores.

## Conclusion

The exploratory nature of the preliminary study discussed here revealed evidence of patterns of individual variation in acquisition of English wh-questions, but also evidence of different patterns of processing in different modes (oral vs. written) and for different types (subject vs. object extraction). The unexpectedly wide range of individual variation in development during immersion and the cross-task variation in WM scores did not provide a robust basis for the hypotheses tested here that WM would correlate with linguistic development during immersion. The theoretical and methodological limitations of the preliminary study are being addressed in a second study, in order to contribute further to our understanding of the L2 user's 'coalition of resources' (Herschensohn, 2000), of the complex interface between input and memory and how WM is involved in the process of second language acquisition.

## Notes

- <sup>1</sup> Thirty participants were originally recruited, but only eleven remained throughout the length of the study, so I only refer to their data here.
- <sup>2</sup> Referring to the greater speed with which Chinese single syllable digits can be spoken in comparison to, say, English.
- <sup>3</sup> Fehringer and Fry (2007) tested story recall on bilingual and near-native L2 speakers of German and English; I am grateful to the authors for permission to adapt their story recall test for this study. However, I am unaware of any research using such a task for Mandarin L1.

## References

- Andrade, J. (Ed.) (2001). *Working memory in perspective*. Hove: Psychology Press.
- Baddeley, A. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–23.
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communications Disorders*, 36, 189–208.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158–73.
- Baddeley, A., & Hitch, G. (1974). Working memory. In G. Bower (Ed.). *The psychology of learning and motivation* (pp. 47–89). New York: Academic Press.
- Brown, G., & Hulme, C. (1992). Cognitive processing and second language processing: The role of short term memory. In R. Harris (Ed.). *Cognitive processing in bilinguals* (pp. 105–21). Amsterdam: Elsevier.
- Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22(1), 77–94.
- Carroll, S. (2001). *Input and Evidence*. Amsterdam: John Benjamins.
- Chincotta, D., & Underwood, G. (1997). Digit span and articulatory suppression: a cross-linguistic comparison. *European Journal of Cognitive Psychology*, 9, 89–96.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42.
- Conway, A., Kane, M., Bunting, M., Hambrick, D., Wilhelm, O., & Engle, R. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, 12, 769–86.
- Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. de Groot, & J. Kroll (Eds.). *Tutorials in bilingualism: psycholinguistic perspectives* (pp. 279–300). Mahwah, NJ: Lawrence Erlbaum.
- Coughlan, A., & Hollows, S. (1985). *The Adult Memory and Information Processing Battery (AMIPB) Test Manual*. Leeds: A K Coughlan.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Mikaye, & P. Shah (Eds.). *Models of working memory* (pp. 62–101). Cambridge: Cambridge University Press.

- Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–84.
- Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–66.
- Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25, 1–18.
- Dornyei, Z. (2003). *Questionnaires in second language research*. Mahwah, NJ: Lawrence Erlbaum.
- Ellis, N. (1994). *Implicit and explicit learning of languages*. London: Academic Press.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305–52.
- Ellis, N., & Sinclair, S. (1996). Working memory in the acquisition of vocabulary and syntax: putting language in good order. *Quarterly Journal of Experimental Psychology*, 49A, 234–50.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford, Oxford University Press.
- Emerson, M., Miyake, A., & Rettinger, D. (1999). Individual differences in integrating and coordinating multiple sources of information. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 1300–21.
- Fehrer, C., & Fry, C. (2007). Frills, furbelows and activated memory: Syntactically optional elements in the spontaneous language production of bilingual speakers. *Language Sciences*, 29, 497–574.
- Fortkamp, M. (1999). Working memory capacity and aspects of L2 speech production. *Communication and Cognition*, 32, 259–95.
- Fry, C. (2002). Language complexity, working memory and social intelligence. Unpublished PhD thesis, Newcastle University, Newcastle, UK.
- Gathercole, S. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 513–43.
- Gathercole, S., & Baddeley, A. (1993). *Working memory and language*. Hillsdale, NJ: Lawrence Erlbaum.
- Gu, Y. (2003). Fine brush and freehand: The vocabulary-learning art of two successful Chinese EFL learners. *TESOL Quarterly*, 37, 73–10.
- Han, Z. (2004). *Fossilisation in adult second language acquisition*. Clevedon: Multilingual Matters.
- Harrington, M., & Sawyer, M. (1992). L2 working memory and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25–38.
- Hawkins, R. (2001). *Second language syntax*. Oxford: Blackwell.
- Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition. *Second Language Research*, 13, 187–226.
- Herschensohn, J. (2000). *The second time around: minimalism and L2 Acquisition*. Amsterdam: John Benjamins.
- Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.
- Johnson, J., & Newport, E. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39, 215–58.
- Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, 102, 199–225.



- Juffs, A., & Rodriguez, G. (2006). Second language sentence processing and working memory in college-educated and low-educated learners of English as a second language. *EuroSLA Conference*, Antalya, Turkey.
- Lado, R. (1965). Memory span as a factor in second language learning. *International Review of Applied Linguistics*, 3, 123–9.
- McDonald, J. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381–401.
- Miyake, A., Carpenter, P., & Just, M. (1994). Working memory constraints on the resolution of lexical ambiguity. *Journal of Memory and Language*, 33, 175–202.
- Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy, & L. Bourne (Eds.). *Foreign language learning* (pp. 339–64). Mahwah, NJ: Lawrence Erlbaum.
- Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory*. Cambridge: Cambridge University Press.
- Mizera, G. (2006). Working memory and L2 oral fluency. (Unpublished PhD thesis), University of Pittsburgh.
- Murphy, V. (1997). The effect of modality on a grammaticality judgement task. *Second Language Research*, 13, 34–65.
- Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30, 287–9.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam/Philadelphia: John Benjamins.
- Pienemann, M. (1998). *Language processing and second-language development: processability theory*. Amsterdam, John Benjamins.
- Sagarra, N. (2000). The longitudinal role of working memory on adult acquisition of L2 grammar. *Dissertation Abstracts International*, A (Humanities and Social Sciences), 61 (5).
- Sagarra, N. (2007). Working Memory and L2 processing of redundant grammatical forms. In Z. Han, E. S. Park, A. Revesz, C. Combs, & J. H. Kim (Eds.). *Understanding second language process* (pp. 133–47). Clevedon: Multilingual Matters.
- Sakas, W., & Fodor, J. (2001). The structural triggers learner. In S. Bertolo (Ed.). *Language acquisition and learnability* (pp. 172–233). Cambridge: Cambridge University Press.
- Schachter, J. (1974). An error in error analysis. *Language Learning*, 24, 205–14.
- Schachter, J., & Yip, V. (1990). Why does anyone object to subject extraction? *Studies in Second Language Acquisition*, 12, 379–92.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–58.
- Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.). *Implicit and explicit learning of languages* (pp. 165–210). London, Academic Press.
- Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behaviour. *Studies in Second Language Acquisition*, 15, 147–63.



- Schwartz, B., Ma, J., & Kim, J. (2008). Island sensitivity in development: L2 adults, L2 children and L2 youths. Third Newcastle University Postgraduate Conference in Theoretical and Applied Linguistics, Newcastle upon Tyne, UK.
- Service, E. (1992). Phonology, working memory, and foreign language learning. *Quarterly Journal of Experimental Psychology A*, 45 (1), 21–50.
- Service, E., Simola, M., Metsanheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology*, 14 (3), 383–408.
- Skehan, P. (1998). *A cognitive approach to learning language*. Oxford: Oxford University Press.
- Smith, E., & Kosslyn, S. (2007). *Cognitive Psychology: Mind and Brain*. New York: Prentice Hall.
- Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4, 105–22.
- VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.
- VanPatten, B. (2005). Processing instruction. In C. Sanz (Ed.), *Mind and Context in Adult Second Language Acquisition* (pp. 267–81). Washington, DC: Georgetown University Press.
- Waters, G., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments and Computers*, 35(4), 550–64.
- White, L. (2003). *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.
- White, L., & Genesee, F. (1996). How native is near native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, 11, 233–65.
- White, L., & Juffs, A. (1998). Constraints on wh-movement in two different contexts of nonnative language acquisition: Competence and processing. In S. Flynn, G. Martohardjono, & W. O'Neil (Eds.), *The generative study of second language acquisition* (pp. 111–29). Mahwah, NJ: Lawrence Erlbaum.
- Wright, C. (2006). Is interlanguage variation in complex CP structures predictable? *Proceedings of the Fourth University of Cambridge–Postgraduate Conference in Language Research*, 277284.
- Yoshimura, Y. (2001). The role of working memory in language aptitude. In X. Bonch-Bruевич, W. Crawford, J. Hellerman, C. Higgins, & H. Nguyen (Eds.), *Second Language Research Forum* (pp. 144–63). Somerville, MA: Cascadilla Press.